

# Statistical Global Modeling of $\beta^-$ -Decay Half-lives Systematics Using Multilayer Feedforward Neural Networks and Support Vector Machines

N. Costiris,<sup>a</sup> E. Mavrommatis,<sup>a</sup> K. A. Gernoth,<sup>b</sup> J. W. Clark,<sup>c</sup> and H. Li<sup>d</sup>

<sup>a</sup>*Physics Department, Section of Nuclear and Particle Physics  
University of Athens, GR-15771 Athens, Greece*

<sup>b</sup>*Institut für Theoretische Physik, Johannes-Kepler-Universität,  
A-4040 Linz, Austria*

*School of Physics and Astronomy, Schuster Building  
The University of Manchester, Manchester, M13 9PL, UK*

<sup>c</sup>*McDonnell Center for the Space Sciences and Department of Physics  
Washington University, St. Louis, MO 63130, USA*

<sup>d</sup>*Dartmouth Medical School, Lebanon, NH 03756, USA*

---

## Abstract

In this work, the beta-decay half-lives problem is dealt as a nonlinear optimization problem, which is resolved in the statistical framework of Machine Learning (LM). Continuing past similar approaches, we have constructed sophisticated Artificial Neural Networks (ANNs) and Support Vector Regression Machines (SVMs) for each class with even-odd character in  $Z$  and  $N$  to global model the systematics of nuclei that decay 100% by the  $\beta^-$ -mode in their ground states. The arising large-scale lifetime calculations generated by both types of machines are discussed and compared with each other, with the available experimental data, with previous results obtained with neural networks, as well as with estimates coming from traditional global nuclear models. Particular attention is paid on the estimates for exotic and halo nuclei and we focus to those nuclides that are involved in the r-process nucleosynthesis. It is found that statistical models based on LM can at least match or even surpass the predictive performance of the best conventional models of  $\beta$ -decay systematics and can complement the latter.

---

\* URL: [www.pythaim.phys.uoa.gr](http://www.pythaim.phys.uoa.gr), E-mail: [pythaim@phys.uoa.gr](mailto:pythaim@phys.uoa.gr)

## 1 Introduction

Reliable, quantitative estimates of  $\beta^-$ -decay halflives of nuclei far from stability are needed by the experimental exploration of the nuclear landscape at existing and future radioactive ion-beam facilities and by ongoing major efforts in astrophysics towards understanding of supernova explosions, and the processes of nucleosynthesis, notably the r-process [1]. In the nuclear chart there are spaces for some 6000 nuclides between the  $\beta$ -stability line and the neutron-drip line. Although,  $\beta^-$ -decay properties have been already measured in terrestrial laboratories for some r-process key nuclei and more will be measured in future facilities, the majority of  $\beta$ -decay rates of the involved neutron-rich nuclei should be estimated from theoretical models. Several approaches of different level of sophistication for determining  $\beta$ -halflives have been proposed and applied over the years. One can mention the more phenomenological treatments based on Gross Theory (*GT*) [2], as well as microscopic treatments that employ the *pn* Quasiparticle Random-Phase Approximation (*pn-QRPA*) (in various versions) [3], [4] or the shell-model [5]. The latest hybrid version of the *RPA* models developed by Möller and coworkers, combines the *pn-QRPA* model with the statistical Gross Theory of the first forbidden decay (*pnQRPA+ffGT*) [6]. There are also some models in which the ground state of the parent nucleus is described by the extended Thomas-Fermi plus Strutinsky integral method, or the Hartree-Fock *BCS* method, or other density functional method (*DF*) and which use the continuum *QRPA* (*CQRPA*) [7]. Recently relativistic *pn-QRPA* (*RQRPA*) models have been applied in the treatment of neutron-rich nuclei in the  $N \sim 50$ ,  $N \sim 82$  and  $Z \sim 28$  and 50 regions [8]. Despite continuing improvements the predictive power of these conventional “theory-thick” models is rather limited for  $\beta^-$ -decay halflives of nuclei that are mainly far from stability, with deviations from experiment of at least an order of magnitude and considerable sensitivity to quantities that are poorly known.

The recent advances in Artificial Intelligence (*AI*) and especially in statistical learning theory or Machine Learning (*LM*), notably Artificial Neural Networks (*ANNs*) and Support Vector Machines (*SVMs*) provide, an alternative opportunity to develop statistical models of observables of different systems that exhibit significant power. These models are “theory-thin” and are driven mainly from the data. For example, in the case of nuclei, any nuclear observable  $X$  can be viewed as a mapping from the proton and neutron numbers  $Z$  and  $N$  ( $((Z,N) \rightarrow X)$ ). In *LM* one attempts to approximate the mapping by using only a subset of the data for  $X$  (training data). *LM*-based models have already been developed for several nuclear properties [9], including atomic masses and ground state spins and parities. In this work, which continues previous studies in statistical modeling of nuclear halflife systematics [10-12], we present global models for the halflives of nuclear ground states that decay

100% by the  $\beta^-$  mode ( $T_\beta$ ) using the *ANN* and *SVM* algorithms aiming to compare the two algorithms and to investigate further the potentiality of the modeling using learning machines. In Section 2 we briefly present the elements of the models. In Section 3 some of the results of the models are presented and evaluated and in Section 4 the conclusion and prospects of the present study are considered.

## 2 The Models

A Learning Machine consists of i) an input interface where, external, input variables to the device in coded form (for ex.  $Z$  and  $N$ ) are fed , ii) a system of intermediate elements or units that process the input, and iii) an output interface where an estimate of the corresponding observable of interest (say the beta half-life  $T_\beta$ ) appears for decoding. Given an adequate body of training data (learning set), a suitable training algorithm is used to adjust the parameters of the machine to produce good performance on this set and good generalization on test examples (test set) absent from the training set. The machine interpolates or extrapolates. *ANNs* are nonlinear computational structures, inspired by biological neural systems, which consist of interconnected group (layers) of artificial neurons (processing units). The connections of the units (weights) determine their function [13]. In particular, feedforward networks (multilayer perceptrons) on which we focus in this work although fully-connected have no lateral and feedback connections and the information flows from the input to the output. *SVMs* are learning systems having a rigorous basis in the statistical learning theory developed by Vapnik and Chervonenkis [14] (VC theory) and belong to the class of Kernel methods (meaning that they implicitly perform a nonlinear mapping of the input data into a high-dimensional feature space). There are similarities as well as differences between *ANNs* and *SVMs*. The differences have mainly to do with the tradeoff between complexity and generalization ability and with the “nature” of their parameters. During *SVM* training a reduced number of training patterns (support vectors) is picked and determines the architecture (optimal number of neurons).

In this work we have constructed four separate *ANN* and *SVM* models that determine the half-lives of the nuclides according to the pairing of  $Z$  and  $N$  (even- $Z$ -even- $N$  (*EE*), even- $Z$ -odd- $N$  (*EO*), odd- $Z$ -even- $N$  (*OE*), and odd- $Z$ -odd- $N$  (*OO*)). We briefly list below the main features of these models and further information on the methodology is found in Ref. [9, 15] and [11, 16] for *ANN* and *SVM* models respectively. The four *ANN* models are fully-connected, multilayer feedforward networks with architecture symbolized by  $[2 - 5 - 5 - 5 - 1|81]$ . The activation function of the neuron-like units is given by hyperbolic-tangent sigmoid function in the intermediate (hidden) layers and a saturated linear function in the output layer. The two input

units encode  $Z$  and  $N$  and the single output unit  $\log_{10} T_\beta$ . The Levenberg-Marquardt backpropagation algorithm has been used to train the network while implementing Bayesian regularization and cross-validation to improve generalization and using the Nguyen-Widrow method for initialization of the network. The four *SVM* models use a *pA*-type Kernel (a composite kernel obeying Mercer’s theorem and constructed linearly by combining the Anova (*A*) kernel and a polynomial (*p*) kernel) with the two *pA* parameters  $p$  and  $d$  equal to 1 and 8 respectively and the control parameter  $C$  equal to  $3.861 \times 10^3$  for all four classes.

The experimental data used in developing our models of  $\beta$ -decay systematics have been taken from the *Nubase2003* evaluation [17]. Restricting attention to those cases in which the ground state of the parent decays 100% through the  $\beta^-$  channel, we form a subset of the beta-decay data denoted by NuSet-A, consisting of 905 nuclides. We also formed a more restricted data set, called NuSet-B, by eliminating from NuSet-A those nuclei having halflife greater than  $10^6$  s. The halflives in this subset range from  $0.15 \times 10^{-2}$  s for  $^{35}\text{Na}$  to  $0.20 \times 10^6$  s for  $^{247}\text{Pu}$ . NuSet-B consists of 838 nuclides (*Overall set*): 672 (80%) (*learning set*) of them have been randomly chosen to train the *ANNs* or to find the support vectors in the case of *SVMs*; of those left 83 (10%) (*validation set*) have been similarly chosen to validate the learning procedure in the case of *ANNs* or to guide the determination of small number of parameter, mainly entering the inner-product Kernel in the case of *SVMs*; the rest 83 ( $\sim 10\%$ ) (*test set*) have been similarly chosen to evaluate the accuracy of the prediction. Considering performance on the above sets we speak of operation in overall, learning, validation and prediction modes. Having excluded the few long-lived examples from NuSet-A, one is then dealing with a more homogeneous collection of nuclides, a property that facilitates the training of network models. Accordingly, we have focused our efforts on NuSet-B, and since the examples still span 9 orders of magnitude in  $T_\beta$ , it is natural to work with the  $\log_{10} T_\beta$ .

### 3 Results And Discussion

The performance of our *ANN* and *SVM* global models is first evaluated in Table 1 by direct comparison with the experimental data using a commonly known statistical metric, namely the *Root Mean Square Error* ( $\sigma_{\text{RMSE}}$ ):

$$\sigma_{\text{RMSE}} = \left[ \frac{1}{N} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]^{1/2}. \quad (1)$$

where  $y_i \equiv \log_{10} T_{\beta,\text{calc}}$  and  $\hat{y}_i \equiv \log_{10} T_{\beta,\text{exp}}$  and  $N$  the total number of nuclides in each case. We mention for comparison the  $\sigma_{\text{RMSE}}$  values of an earlier *ANN* model [12] which equal to 1.08 and 1.82 for total learning and test sets, respectively, as well as those of the *ANN* model developed recently by means of the whole basis [10], which equal to 0.53, 0.60 and 0.65 for the total learning, validation and test sets respectively. In Table 2, a further assessment is made by tabulating the performance measures of Möller and collaborators [6]:  $M$ , its standard deviation  $\sigma_M$  and  $\Sigma$ . These measures are defined as follows in terms of the variable  $r_i = y_i/\hat{y}_i$ :

$$M = \frac{1}{N} \sum_{i=1}^N r_i, \quad \sigma_M = \left[ \frac{1}{N} \sum_{i=1}^N (r_i - M)^2 \right]^{1/2}, \quad \Sigma = \left[ \frac{1}{N} \sum_{i=1}^N r_i^2 \right]^{1/2}. \quad (2)$$

Superior models should have  $M$ ,  $\sigma_M$  and  $\Sigma$  near zero. A comparison follows with two recent global theory-thick models of the above collaboration, namely the *FRDM+pnQRPA* and the *pnQRPA+ffGT* models. [6]. Finally, in Fig. 1, halflives of  $\beta^-$ -decaying nuclides that are found near or on a typical r-process path with neutron separation energy below 3 MeV derived by means of the present *LM*-models are compared with those from *pnQRPA+ffGT* calculation and a calculation by Pfeiffer and coworkers [18] (labeled GT\*) based on the early Gross Theory (GT) of Takahashi et al. [2a] with updated mass values.

From the above comparison one can conclude that *LM*-based models give similar results, which are close to experimental data. In Fig. 1, the results given by the *SVM* model are almost equal with the experimental values. This occurs because, unlike neuron network, where the connections between neurons are random values, SVMs use exclusively only almost all training patterns as “nodes”, a fact that may lead to better interpolation but often worse extrapolation. Furthermore, the comparison of the results derived by the *LM*-models with those of the models presented in Table 2 as well with others [15,16] lead to the conclusion that the former perform equally or better than the later. This is partially ascribed to the larger number of parameters of the *LM*-based models. Regarding the performance of the present *ANN* and *SVM* models with respect to that of previous *LM*-based models the consideration of different quality measures slightly favors the present ones. However, a more detailed analysis (see Ref. [15]) shows that the subdivision of the data into four  $(Z,N)$  parity classes can lead to spurious fluctuations in the prediction of lifetimes for nuclides of isotopic and isotonic chains. This favors the use of the *ANN* model developed recently by means of the whole basis [10].

Table 1

Root-mean-square errors  $\sigma$  ( $\equiv \sigma_{\text{RMSE}}$ ) (Eq. 1) for learning, validation and test sets, achieved by the current *ANN* and *SVM* models (developed for even-even, even-odd, odd-even and odd-odd classes, in  $Z$  and  $N$ ) of  $\beta^-$ -decay halflives (with a cutoff at  $10^6 s$ ).

Class	<i>ANN</i> models						<i>SVM</i> models					
	Lear. Set		Val. Set		Test Set		Lear. Set		Val. Set		Test Set	
	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$	$N$	$\sigma$
EE	131	0.36	16	0.41	16	0.62	131	0.55	16	0.57	16	0.62
EO	179	0.38	22	0.44	22	0.39	179	0.41	22	0.42	22	0.51
OE	172	0.44	21	0.46	21	0.53	172	0.41	21	0.47	21	0.47
OO	190	0.52	24	0.42	24	0.33	190	0.52	24	0.40	24	0.52
<b>Total</b>	<b>672</b>	<b>0.41</b>	<b>83</b>	<b>0.44</b>	<b>83</b>	<b>0.51</b>	<b>672</b>	<b>0.47</b>	<b>83</b>	<b>0.46</b>	<b>83</b>	<b>0.53</b>

Table 2

Quality indices  $M$ ,  $\sigma_M$  and  $\Sigma$  (Eq. 2) for the present *ANN* models in Overall (a) and Prediction (b) Modes and for the *FRDM*+*pnQRPA* and *pnQRPA*+*ffGT* models of Ref. [6]. The number  $n$  stands for the nuclides with experimental halflives below the prescribed limit.

$\mathbf{T}_{\beta,exp}$	<b>(a) <i>ANNs</i> - Overall Mode</b>				<b>(b) <i>ANNs</i> - Prediction Mode</b>			
( $s$ )	$n$	$M$	$\sigma_M$	$\Sigma$	$n$	$M$	$\sigma_M$	$\Sigma$
$< 1$	252	0.02	0.30	0.30	8	-0.19	0.61	0.64
$< 10$	396	0.02	0.35	0.35	28	-0.16	0.57	0.60
$< 100$	529	0.03	0.36	0.36	54	-0.05	0.51	0.52
$< 1000$	653	0.05	0.39	0.39	77	0.03	0.49	0.49
$< 10^6$	838	0.00	0.45	0.45	83	0.03	0.52	0.53
$\mathbf{T}_{\beta,exp}$	<b>(c) <i>FRDM</i> + <i>pnQRPA</i> [6]</b>				<b>(d) <i>pnQRPA</i> + <i>ffGT</i> [6]</b>			
( $s$ )	$n$	$M$	$\sigma_M$	$\Sigma$	$n$	$M$	$\sigma_M$	$\Sigma$
$< 1$	184	0.03	0.57	0.57	184	-0.08	0.48	0.49
$< 10$	306	0.14	0.77	0.78	306	-0.03	0.55	0.55
$< 100$	431	0.19	0.94	0.96	431	-0.04	0.61	0.61
$< 1000$	546	0.34	1.28	1.33	546	-0.04	0.68	0.68
$< 10^6$	—	—	—	—	—	—	—	—

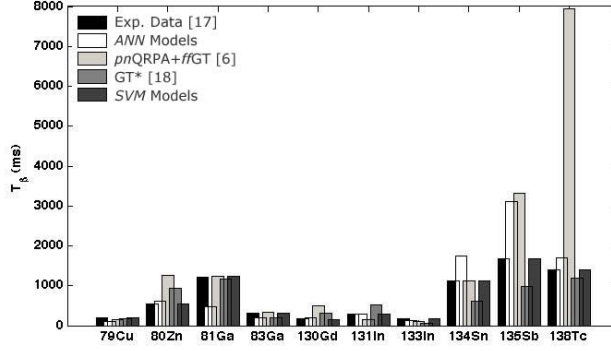


Fig. 1. Halfives for  $\beta^-$ -decaying nuclides that are found near or on a typical r-process path with the neutron separation energy lower or equal to 3 MeV derived by means of the present *ANN* and *SVM* models are compared with the experimental data and those from *pnQRPA+ffGT* [6] and *GT\** [18] calculations.

## 4 Conclusion and Prospects

In this work, the beta-decay halfives problem is dealt as a nonlinear optimization problem, which is resolved in the statistical framework of Machine Learning using Artificial Neural Networks (*ANNs*) and Support Vector Regression Machines (*SVMs*). It seems that both *ANNs* and *SVMs* demonstrate similar performance and that our statistical large-scale calculations can match or even surpass the predictive performance of the best conventional global calculations outside the stable valley. Moreover, this way of confrontation of the beta-halfives problem could give an estimation of the degree that the nucleonic numbers determine the beta-decay systematics of a nuclear system. Accordingly, we plan further studies of the systematics of beta decay using *LMs*, with the object of continued enhancement of their predictive power and of possible gaining of some new physical insight.

## 5 Acknowledgments

This research has been supported in part by the U.S. National Science Foundation under Grant No. PHY-0140316 and by the University of Athens under Grant No.70/4/3309.

## References

- [1] NUPECC Long Range Plan 2004: Perspectives for Nuclear Physics Research in Europe in the Coming Decade and Beyond (April 2004); Opportunities

in Nuclear Science: The Frontiers of Nuclear Science: A Long Range Plan (DOE/NSF, 2007).

- [2] (a) K. Takahashi, M. Yamada and T. Kondoh, At. Data Nucl. Data Tables **12** (1973) 101; (b) H. Nakata, T. Tachibana and M. Yamada, Nucl. Phys. **A625** (1997) 521.
- [3] For the latest models of the Klapdor group see: H. Homma, M. Bender, M. Hirsch, K. Muto, H. V. Klapdor and T. Oda, Phys. Rev. **C54**:6 (1996) 2972; J. U. Nabi and H. V. Klapdor, At. Data Nucl. Data Tables **71** (1999) 149; *ibid* **88** (2004) 237.
- [4] P. Möller, J. R. Nix and K.-L. Kratz, At. Data Nucl. Data Tables **66** (1997) 131 and references therein.
- [5] For the latest calculations see: H. Grawe, K. Langanke and G. Martinez - Pinedo, Rep. Progr. Phys. **70** (2007) 1525.
- [6] P. Möller, B. Pfeiffer and K.-L. Kratz, Phys. Rev. **C67**:5 (2003) 055802.
- [7] See for ex.: J. Engel et al., Phys. Rev. **C60** (1999) 014302; I. N. Borzov, Nucl. Phys. **A777** (2006) 645.
- [8] T. Niksic, T. Marketin, D. Vretenar, N. Paar and P. Ring, Phys. Rev. **C71**:1 (2005) 014308; T. Marketin, D. Vretenar and P. Ring, Phys. Rev. **C75** (2007) 024304.
- [9] J. W. Clark, in Scientific Applications of Neural Nets, J. W. Clark, T. Lindenau and M. L. Ristig (eds) (Springer-Verlag, Berlin, 1999) p1; K. A. Gernoth, *ibid*, p. 139.
- [10] N. J. Costiris, E. Mavrommatis, K. A. Gernoth and J. W. Clark, in Proceedings of the 16th Hellenic Symposium of the Hellenic Nuclear Physics Society, (Athens, 2006) 210 (nucl-th/0701096).
- [11] J. W. Clark and H. Li, Int.J.Mod.Phys. **B20** (2006) 5015 (nucl-th/0603037).
- [12] E. Mavrommatis, A. Dakos, K. A. Gernoth and J. W. Clark, in Condensed Matter Theories, Vol. 13, edited by J. da Providencia and F. B. Malik (Nova Sciences Publishers, Commack, NY, 1998) 423.
- [13] S. Haykin, *Neural Networks: A Comprehensive Foundation* (McMillan, N.Y.) 1993.
- [14] V. Vapnik, *The Nature of Statistical Learning Theory* (Springer, N.Y.) 1995.
- [15] N. Costiris, *Diploma Thesis* (University of Athens, Greece) 2006.
- [16] H. Li, *Ph-D Thesis* (Washington University, USA) 2008.
- [17] G. Audi, O. Bersillon, J. Blachot and A. H. Wapstra, Nucl. Phys. **A729** (2003) 3.
- [18] B. Pfeiffer, K. L. Kratz and P. Möller, (Institut Für Kernchemie, Internal Report) 2003.